Filling the gaps: Using synthetic aperture radar data to reconstruct multispectral images

Expert: Kristian Bodolai (Space Intelligence)

Session: TBD

Satellite data is great. It allows us to see what's going on on the earth's surface frequently, allowing a multitude of applications like forest monitoring, change detection, disaster relief…

It mainly comes in two flavours, active sensors and passive sensors. Passive sensors measure sunlight bouncing off the earth (very much like a regular camera, but normally looking beyond RGB), while active sensors create their own radiation and measure it back (like lidar or radar).

Moreover, there is a massive collection of satellite data available for free, such as the Sentinel and Landsat missions, among many others. This allows governments, NGOs and private companies, such as Space Intelligence, to create real impact using this high quality public data.

One outstanding problem for remote sensing of the earth is clouds. 67% of Earth's surface is covered by clouds on average. Luckily, radiation in the radar portion of the spectrum can pass through clouds, but many of the frequencies used in remote sensing can't.

Standard practice is using a classifier or some bands to detect clouds, mask them out of the scenes, and mosaicking several scenes so that the gaps get filled once we add enough images. This usually works very well, but some areas, specially in the tropic, are so cloudy that you need to combine a year's worth of data to get a relatively clean image. Losing a lot of information in the time-series in the process.

The aim of the project is to use data form the European Space Agency missions Sentinel-1 (S1 - synthetic aperture radar data) and Sentinel-2 (S2 - Multispectral, 13 bands, including RGB) to answer one, more or a variant of the following questions:

- Can S1 be used to reconstruct a S2 vegetation index?
- Can S1 be used to reconstruct S2 RGB data?
- Can S1 be used to reconstruct all S2 bands? Is there any relationship between band wavelength and error?
- What are the uncertainties associated with this? Are there any land cover types where this works specially good or specially bad?
- How else can we frame the problem to enhance information retrieval?

Keep in mind that this is a hard problem. Radar sensors and multispectral sensors ultimately 'see' very different things, so there may be an irreducible error associated with this.

You can read more about ESA's sentinel missions here and a bit about remote sensing in general here. Dealing with satellite images requires handling of geospatial data (images + coordinates), this can be easily done using GDAL or rasterio (in Python, but most programming languages have some version of this).